ACCELERATE YOUR ENTERPRISE DATA MIGRATION WITH INFOSYS DATA WIZARD

Infosys
topaz

databricks

Infosys®
Navigate your next

Enterprise data migration involves moving an organization's data from one platform to another. While the premise sounds simple, the process is traditionally complex. Data migration is one of the most critical projects any enterprise can undertake, where one misstep can cause several problems—from security issues to operational dysfunction. Other challenges could include a lack of scalability, long project timelines, data quality issues, and high costs. Studies reveal that 39% of cloud spending goes to data management, with 82% of data professionals unable to control or forecast data costs.

As data volumes grow and systems become more complex, there is a need for planning and designing the migration process through intelligent solutions that ensure secure, efficient, and seamless enterprise migration. Infosys Data Wizard for Databricks with Infosys Cognitive Code Converter helps accelerate and de-risk enterprise data migrations by leveraging a metadata-driven codeless framework. Powered by generative artificial intelligence (GenAI) frameworks based on cutting-edge large language models (LLMs), it facilitates existing data and code analyses, explanation, code conversion, and unit testing. Its intuitive, user-friendly interface enables end-to-end migration automation with modular, optimized processes and extensive validation.

Our solution delivers up to three times faster migration speeds and reduces costs by up to 30%. It offers a demonstrated 30% acceleration in operationalizing AI while slashing cloud consumption costs by 20%. Infosys Data Wizard also supports a wide range of legacy data platforms, including Teradata, Netezza, Oracle, Informatica, mainframes, and Hadoop distributions such as Cloudera, enabling migration to Databricks Delta and Databricks SQL (DBSQL).

Enterprise Challenges

Many organizations struggle with manual processes and subpar data quality that delay data migrations to modern digital systems. In the long run, these challenges lead to substantial costs and significant operational inefficiencies.

Migrating from legacy systems such as Teradata and Hadoop to more modern platforms like the Databricks Data Intelligence Platform is crucial for enterprises looking to modernize their data infrastructure. However, there are barriers to migration that impact the entire lifecycle from configuration to post-migration validation.

These include:

1. **Business challenges**
   - Legacy systems that are difficult to integrate with modern platforms
   - Data silos that hinder integration and collaboration
   - Diverse workloads and a complex IT stack that require significant effort to manage

   - Operational challenges from monolithic architecture
   - High modernization, maintenance, and production costs

2. **Configuration challenges**
   - Manual and unsecured connectivity between source and target systems for data pipelines
   - Inefficient user and credential management processes
   - Insecure connections and complex access mechanisms

3. **Accurate inventory discovery**
   - Reliance on complex script-driven inventory collection
   - Use of multiple scripts to extract varying levels of detail

4. **Historical and ongoing delta data and code analyses**
   - Poor data and code quality
   - Manual processes for detecting data patterns and analyzing code
   - Limited data access and skill-related barriers in data handling
   - Manual tracking and monitoring of migration progress

5. **Migration execution challenges**
   - Extensive manual migration and intervention requiring significant effort, often resulting in delays
   - Absence of timely notifications or updates on status

6. **Post-migration validation issues**
   - Exhaustive and error-prone manual validation and testing processes resulting in multiple iterations of migrations, data reloads, and code refactoring
   - Gaps between different stages of migration, creating challenges in regression testing and robust end-to-end validation

Streamlining Data Migration with Infosys Data Wizard

Migrating from legacy systems is often a time-consuming and resource-intensive process. Enterprises that take the traditional route rely on manual inventory assessments, fragmented tools, and long-drawn code redevelopment. Such an approach could lead to inefficiencies, higher costs, and risk of errors. Infosys, a Global Elite SI Partner of Databricks, helps enterprises overcome these challenges with Infosys Data Wizard for Databricks and Infosys Cognitive Code Converter.

Leveraging the open Databricks Data Intelligence Platform, which includes components such as Unity Catalog, Mosaic AI, and Databricks SQL, our solution helps accelerate migrations while minimizing manual effort, errors, and expenses. Infosys specializes in data warehouse and data lake migration, as well as modernization using the asset suite of the Infosys Autonomous Data and AI Estate Solution, which leverages AI/machine learning (ML) and GenAI to accelerate value delivery.

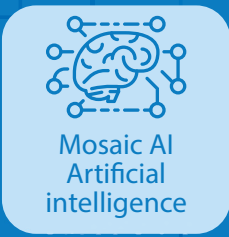# Databricks Data Intelligence Platform

100% serverless

| Disaster recovery | Cost controls | Enterprise security |

| Mosaic AI Artificial intelligence | Databricks SQL Data warehousing | Workflows/DLT Ingest, ETL, streaming | AI/BI Business intelligence |

Lakehouse

Unity Catalog

DELTA LAKE   **ICEBERG**   **Parquet**

Some of the key features of our solution include:

Automated inventory collection and analysis

Infosys Data Wizard leverages traditional and AI-led techniques to gather inventory-related information from data, code, and business intelligence (BI) layers, driving comprehensive inventory analysis and rationalization.

AI-driven framework for accelerated data migration

Our solution accelerates the migration of data warehouses and data lakes to the Databricks Data Intelligence Platform using data extraction and transfer mechanisms. Our AI-driven framework ensures secure data migration from legacy systems to Databricks while optimizing migration code and BI transformation for seamless integration.

Comprehensive certification

Infosys Data Wizard automates the validation of migrated data, code, and BI reports to certify a successful migration from legacy platforms to the Databricks Data Intelligence Platform.

De-risk your Data Migration with Infosys Data Wizard

Infosys Data Wizard offers comprehensive optimization by identifying and eliminating inefficiencies. We offer tailored migration plans, strategies, and blueprints to ensure the best price-to-performance ratio and the lowest total cost of ownership

(TCO). With AI-driven code transformation, our solution automates code migration for enhanced efficiency and quality.

Our user-friendly migration framework simplifies data and code migration, supported by predictive workload optimization. Infosys Data Wizard also facilitates efficient migration of historical data through direct connect and file-based modes, ensuring continuous data synchronization with incremental data migration.

We provide intelligent data-driven experiences through automated schema mapping and DDL generation for seamless transitions. Infosys Data Wizard offers a streamlined, reliable, and cost-effective migration experience.

Datawarehouse Transformation Success Stories

Infosys Data Wizard seamlessly integrates advanced tools, AI-driven capabilities, and automation to accelerate data migration and transformation for enterprises.

Case study 1: Accelerating agility and efficiency for a leading American manufacturer

A large American manufacturing corporation transformed its data ecosystem by migrating its data and applications from a legacy Netezza Enterprise Data Warehouse (EDW), Hortonworks Data Platform (HDP) data lake, and Informatica extract-transform-load (ETL) environment to Databricks on AWS.

The client leveraged DevOps with containerization—a high-performance serverless platform—and integrated new data sources. By modernizing their enterprise data lake with a focus on scalability, cost optimization, and performance, they significantly improved business agility and reduced operational expenses.

Key outcomes: The migration delivered a 75% improvement in time to market, a 55% cost reduction in data processing, and a 50% increase in efficiency. Incremental data processing of 200GB was reduced from 6 hours to 3 hours, leading to operational savings and greater business agility.

Case study 2: Revolutionizing customer insights for a beverages multinational

A leading multinational beverages chain modernized its legacy customer insights application by moving from Oracle Exadata to Azure Databricks, utilizing Databricks' governance solution, Unity Catalog. This transformation established a cloud-enabled, metadata-driven customer data analytics platform that was adaptable to evolving standards and guidelines. The self-service data management capabilities empowered a lean engineering team to efficiently manage over six petabytes (6PB) of data, more than 1000 engineered datasets, and over 450 data pipelines.

Key outcomes: The transformation resulted in a 40% reduction in costs and empowered analysts to onboard and manage data products independently, significantly improving business agility and operational efficiency.

Case study 3: Transforming data operations for a sports goods major

A leading producer of sports goods in the US migrated from Teradata to Databricks, unlocking substantial business benefits. They migrated to Delta Lake, modernizing complex procedures with efficient PySpark jobs, centralizing metadata and user management, as well as implementing secure data sharing with Delta Sharing. By establishing a robust data quality framework, the client achieved significant gains. The company also utilized Databricks' governance solution, Unity Catalog, providing access controls, auditing and monitoring

Key outcomes: This modernization led to a threefold increase in processing volume, over US $350K in annual compute cost savings and a performance boost with CPU utilization rising from 11% to 50%. Platform stability improved with a 50% reduction in production support incidents, saving an additional US $250K annually. The migration improved data quality and performance monitoring capabilities, ensuring data integrity and better alignment with business needs.

Infosys Topaz is an AI-first set of services, solutions and platforms using generative AI technologies. It amplifies the potential of humans, enterprises, and communities to create value. With 12,000+ AI assets, 150+ pre-trained AI models, 10+ AI platforms steered by AI-first specialists and data strategists, and a 'responsible by design' approach, Infosys Topaz helps enterprises accelerate growth, unlock efficiencies at scale and build connected ecosystems. Connect with us at infosystopaz@infosys.com

For more information, contact askus@infosys.com

Infosys®
Navigate your next